Explainable Artificial Intelligence (XAI) Models for Yield Prediction and Disease Forecasting in *Zea mays* L.

Dr. Carlos Ali [1], Dr. Lucas Minh Anh [2*], Dr. Rafael Van Nam [3], Dr. Gabriel Lan [4]

[1-2] Student, College of Agricultural Biotechnology, Vietnam Agricultural Innovation University, Hanoi, Vietnam
[3] Professor, College of Agricultural Biotechnology, Vietnam Agricultural Innovation University, Hanoi, Vietnam
[4] Teacher, College of Agricultural Biotechnology, Vietnam Agricultural Innovation University, Hanoi, Vietnam

*Corresponding author: **Dr. Lucas Minh Anh**

Received: 21 Nov 2025
Revised: 11 Dec 2025
Accepted: 26 Dec 2025
Published: 22 Jan 2026

E-ISSN: 2989-5340
DOI: https://doi.org/10.54660/jafi.2026.6.1.11-19

ABSTRACT

Background: Maize (*Zea mays* L.) is a cereal crop of strategic importance for global food, feed and bioenergy security, but its productivity is highly susceptible to climatic variability, soil heterogeneity and recurrent foliar and stalk diseases. Traditional machine learning (ML) methods for yield forecasting and disease prediction have demonstrated high predictive accuracy, but remain largely "black bo x" methods, which has limited their uptake by agronomists, extension officers and smallholder farmers.

Objective: In this work, we propose and evaluate an Explainable Artificial Intelligence (XAI) framework that integrates multi-source agronomic data with interpretable machine learning and deep learning architectures to predict maize yield and disease risk, while providing transparent and actionable explanations.

Methods: Data were collected from multi environment field trials with weather data, soil physicochemical properties, vegetation indices derived from unmanned aerial vehicle (UAV) and satellite data, Internet of Things (IoT) sensor data streams and ground truth disease and yield observations. Ensemble tree-based models, convolutional neural networks, and hybrid models were applied for benchmarking. Shapley Additive Explanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) were used to interpret feature contributions.

Results: The gradient-boosted ensemble with SHAP interpretation achieved the best yield prediction accuracy ($R^2 = 0.93$; RMSE = 0.41 t ha^{-1}) and disease forecasting performance (F1-score = 0.91), outperforming traditional deep learning baselines and still being interpretable. The most important predictors across environments were canopy temperature, normalized difference vegetation index (NDVI), relative humidity and nitrogen status.

Significance: The XAI framework significantly increased the model transparency without compromising the predictive performance, producing decision-support recommendations that can be directly linked to the underlying agro-environmental drivers.

Conclusion: Explainable Artificial Intelligence offers a scientifically sound and operationally reliable route towards climate-smart, precision maize production, bridging the gap between algorithmic sophistication and field-level decision making.

Keywords: Precision agriculture, Deep learning, SHAP interpretability, Crop disease modeling, Remote sensing indices, Internet of Things sensing, Decision-support systems, Climate-smart agriculture

1. Introduction

Maize is among the three most cultivated cereals in the world, owing to its multitude of uses as food, livestock feed, and industrial feedstock across various agro-ecologies (Erenstein *et al.*, 2022) [1]. The growing population trend, as well as the changing diets and expanding needs for bioenergy, bring rising demands for maize (Erenstein *et al.*, 2022) [1] (Shiferaw *et al.*, 2011) [2]. Nevertheless, maize production is constantly affected by a number of biotic and abiotic factors, including irregularly distributed rainfall, high temperatures during critical reproduction periods, nutrient variability of the soil and challenges including major developments of crop diseases caused by northern corn leaf blight, gray leaf spot, common rust and *Fusarium* stalk rot— these stresses affect maize productivity (Shiferaw *et al.*, 2011) [2] (Savary *et al.*, 2019) [3]. The nonlinear and site-specific nature of these interactions makes crop yield and disease forecasts very difficult to make using solely conventional agronomic or statistical models (Van Klompenburg *et al.*, 2020) [4] (Liakos *et al.*, 2018) [6].

Artificial Intelligence (AI) has proven to be an effective tool in solving this problem. Machine Learning (ML) and Deep Learning (DL) technologies are increasingly being used for modeling complex relationships between environmental factors, the condition of crops, and the resulting crop productivity (Liakos *et al.*, 2018) [6] (Khaki and Wang, 2019) [5] (Kamilaris and Prenafeta-Boldú, 2018) [23]. It is noteworthy that several algorithms such as random forest, gradient boosting, convolutional neural networks (CNN), and recurrent neural networks (RNN) show an ability to provide yield or disease classification predictions regardless of the cropping systems used (Khaki and Wang, 2019) [5] (Kamilaris and Prenafeta-Boldú, 2018) [23] (Sharma *et al.*, 2021) [25].

However, a significant problem with modern AI tools is that their operation is still not transparent to scientists and practitioners in other related areas (Adadi and Berrada, 2018) [7] (Arrieta *et al*., 2020) [8]. This means that low transparency of the modern prediction methods influences the level of confidence of farmers and agronomists working in the field, who need not only to obtain reliable predictions but also understandable and efficient explanations (Ribeiro *et al*., 2016) [19] (Lundberg and Lee, 2017) [9] (Molnar, 2022) [20].

Explainable Artificial Intelligence (XAI) was developed as a methodology to address this interpretability gap (Adadi and Berrada, 2018) [7] (Arrieta *et al*., 2020) [8]. The term "XAI" describes a family of related techniques including Shapley Additive Explanations (SHAP), Local Interpretable Model-Agnostic Explanations (LIME), partial dependence reasoning, and attention visualization, among others (Arrieta *et al*., 2020) [8] (Lundberg and Lee, 2017) [9] (Ribeiro *et al*., 2016) [19]. All these techniques help assess and communicate how the input features affect the model predictions at either the global level (i.e., over the entire dataset) or the local level (i.e., the specific prediction being made) (Lundberg and Lee, 2017) [9] (Ribeiro *et al*., 2016) [19] (Molnar, 2022) [20]. In an agricultural context, XAI allows researchers to check that the predictions made by the model are based on valid agronomical relationships rather than misleading correlations (Ryo, 2022) [16] (Molnar, 2022) [20].

Advancements in remote sensing technology and the Internet of Things (IoT) provide more information and lead to a more precise analysis of crops through the use of unmanned aerial vehicles (UAVs) provided with RGB, multispectral, hyperspectral, and thermal sensors allowing canopy reflectance and temperature acquisition from the fields as frequently as needed and satellite technology allowing to produce some vegetation indices, including Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) (Weiss *et al*., 2020) [10] (Maes and Steppe, 2019) [11]. Additionally, data from the ground IoT sensors network allow the continuous monitoring of soil moisture, soil temperature, and micrometeorological parameters, thus allowing monitoring the conditions of crop growing almost in real-time (Maes and Steppe, 2019) [11] (Talavera *et al*., 2017) [12]. All this information combined with an artificial intelligence analytical compound enables the creation of decision support systems capable of giving early signals for possible appearance of diseases, nutrient deficiencies, and pressure from water though available data. (Talavera *et al*., 2017) [12] (Chlingaryan *et al*., 2018) [13].

An increasing number of research studies have focused on ML- and DL-based approaches for predicting maize yields with the aid of various climatic, soil, and remote sensing variables, resulting in mostly positive predictions when compared with traditional approaches based on regression (Chlingaryan *et al*., 2018) [13] (Mochida *et al*., 2019) [14]. In a similar fashion, some studies have used hybrid and convolutional techniques to recognize maize leaf diseases based on digital images with an appropriate level of accuracy (Mochida *et al*., 2019) [14] (Mohanty *et al*., 2016) [15]. However, it seems that relatively few studies have utilized explainability frameworks for yield prediction and disease recognition in a holistic framework that integrates various data sources and has even fewer studies examined the significance of interpreting results from XAI in terms of their influence on decision-making support at the field level (Mohanty *et al*., 2016) [15] (Ryo, 2022) [16]. It has to be noted that this is a significant drawback, since the transfer of results from AI to useful advice for farmers depends on the transparency of the models used (Ryo, 2022) [16] (Sishodia *et al*., 2020) [17].

This research study fills in the gap by designing and assessing an integrated XAI system for the purpose of maize yield estimation and disease prognosis employing multi-environment field data, UAV and satellite remote sensing, IoT sensing network as well as standard agronomic practices. The aims of the project were: (I) to measure the ability of various ML and DL frameworks for maize yield forecasting and disease classification; (II) to use SHAP and LIME methods for identification of factors affecting the validity of outputs; (III) to compare explainable and traditional black box frameworks based on their transparency, validity, and efficiency of computations.

2. Materials and Methods

2.1. Study Area

The study was conducted across three agro-climatically distinct maize-growing environments representative of subtropical, semi-arid, and temperate production systems, selected to capture a broad range of thermal regimes, rainfall patterns, and soil types. Site-level characteristics, including geographic coordinates, elevation, mean seasonal temperature and rainfall, and dominant soil classification, are summarized in Table 1. The inclusion of contrasting environments was intended to improve the generalizability of the resulting predictive models under variable climatic conditions, consistent with recommended practice for multi-environment crop modeling trials (Sishodia *et al*., 2020) [17] (Van Ittersum *et al*., 2013) [18].

2.2. Plant Materials

Three commercially released, medium- to long-duration maize hybrids with differential disease susceptibility ratings were selected to introduce genotypic variability into the dataset. Field trials were established using a randomized complete block design with three replications per environment. Agronomic management, including sowing density, fertilization regime, and irrigation scheduling, followed regionally recommended practices, with details summarized in Table 2. Disease pressure was allowed to develop under natural inoculum conditions, supplemented by controlled inoculation in designated sub-plots to ensure adequate representation of disease severity classes.

2.3. Data Collection

A multi-source data acquisition strategy was employed, integrating ground-based agronomic observations with proximal and remote sensing platforms and IoT sensor networks (Weiss *et al*., 2020) [10] (Talavera *et al*., 2017) [12]. Meteorological variables (air temperature, relative humidity, rainfall, solar radiation, and wind speed) were recorded at automated weather stations positioned within each trial site. Soil physicochemical properties, including texture, pH, organic carbon, and macronutrient status, were determined from composite samples collected prior to sowing. Crop growth stage, plant height, leaf area index, and chlorophyll content were recorded at key phenological intervals. UAV-based RGB, multispectral, and thermal imagery were acquired at approximately ten-day intervals throughout the growing season, while satellite-derived multispectral reflectance provided complementary regional coverage (Weiss *et al*., 2020) [10] (Maes and Steppe, 2019) [11]. Hyperspectral imaging was employed on a subset of plots to enhance early-stage disease detection sensitivity (Wang *et al*., 2021) [21]. Disease incidence and severity were scored using standardized visual rating scales at multiple growth stages, and final grain yield was determined by plot-level harvest and adjusted to standard moisture content. IoT-based soil moisture and canopy temperature sensors provided continuous, highfrequency environmental monitoring (Talavera *et al*., 2017) [12]. The categories, sources, and temporal resolution of all datasets are summarized in Table 3, and the derived feature set used for model development is presented in Table 4.

2.4. Data Preprocessing
Prior to model development, all datasets underwent a structured preprocessing pipeline. Data cleaning procedures identified and addressed sensor anomalies, outliers, and inconsistent records. Missing values arising from sensor downtime or cloud-obscured satellite imagery were addressed using context-appropriate imputation strategies, including temporal interpolation for continuous sensor streams and multivariate imputation for sporadically missing agronomic observations. Feature engineering derived composite indices from raw spectral bands, including NDVI, EVI, the Green Normalized Difference Vegetation Index (GNDVI), and canopy temperature depression, alongside cumulative growing degree days and water deficit indices. All continuous features were normalized to a common scale to facilitate model convergence, and categorical variables were appropriately encoded. Feature selection combined correlation-based filtering with embedded importance ranking to reduce dimensionality and mitigate multicollinearity prior to model training (Chlingaryan *et al*., 2018) [13] (Van Klompenburg *et al*., 2020) [4].

2.5. Explainable AI Framework
A tiered modeling strategy was adopted, encompassing conventional machine learning algorithms (multiple linear regression, support vector regression, random forest, and gradient boosting), deep learning architectures (multilayer perceptron, one-dimensional convolutional neural network, and long short-term memory networks for temporal sequences), and ensemble hybrid models combining tree-based and neural predictors. Model architectures, hyperparameter search strategies, and training configurations were selected based on established best practice for agricultural predictive modeling and are described conceptually without reference to specific software implementation (Van Ittersum *et al*., 2013) [18] (Ribeiro *et al*., 2016) [19]. Explainability was incorporated through two complementary approaches. SHAP was applied to quantify the marginal contribution of each input feature to individual and aggregate model predictions, providing both global feature importance rankings and local, instance-level explanations grounded in cooperative game theory (Lundberg and Lee, 2017) [9] (Molnar, 2022) [20]. LIME was applied as a complementary local surrogate-modeling approach, approximating the behavior of the complex model in the neighborhood of individual predictions using interpretable linear surrogates (Molnar, 2022) [20] (Wang *et al*., 2021) [21]. Together, these techniques enabled cross-validation of feature attribution patterns and enhanced confidence in the biological plausibility of model reasoning (Molnar, 2022) [20] (Ryo, 2022) [16]. A decision-support module translated ranked feature attributions and predicted outcomes into categorical risk levels and management recommendations, described further in Section 2.6.

2.6. Prediction Framework
Two primary prediction tasks were addressed: (i) continuous prediction of final grain yield (t ha^{-1}) as a function of environmental, soil, and remote sensing covariates measured up to the mid-reproductive growth stage; and (ii) classification of disease risk into four ordinal severity categories (negligible, low, moderate, high) based on environmental conducive-condition indices, canopy spectral anomalies, and early symptom detection features. A composite stress index integrating water deficit, thermal stress, and nutrient status indicators was additionally derived to support generalized crop stress prediction. Early warning thresholds were established by calibrating predicted disease probability and stress index outputs against historical outbreak records, enabling lead-time risk alerts for use within the decision-support framework.

2.7. Statistical Analysis
Analysis of variance (ANOVA) was used to evaluate the effects of environment, genotype, and their interaction on yield and disease severity. Pearson correlation analysis characterized bivariate relationships among candidate predictors, while Principal Component Analysis (PCA) was applied to explore multivariate structure within the remote sensing and sensor feature space and to support dimensionality reduction. Regression-based diagnostics accompanied all continuous prediction models. Model validation employed five-fold cross-validation stratified by environment, with an independent holdout test set reserved for final performance evaluation. Predictive performance for continuous outcomes was assessed using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE), while classification performance for disease forecasting was evaluated using accuracy, precision, recall, and F1-score. All statistical comparisons were considered significant at $p < 0.05$.

3. Results
3.1. Data Characteristics and Environmental Variability
The multi-environment dataset captured substantial variability in temperature, rainfall, and soil properties across the three trial sites (Table 1), providing a robust basis for model generalization (Van Ittersum *et al*., 2013) [18] (Sishodia *et al*., 2020) [17]. ANOVA indicated significant effects of environment, genotype, and their interaction on both grain yield and disease severity ($p < 0.05$), confirming that environmental heterogeneity was a major driver of phenotypic variability and justifying the multi-source, environment-aware modeling approach adopted in this study (Van Ittersum *et al*., 2013) [18].

3.2. Machine Learning and Deep Learning Model Performance
Comparative evaluation of the seven candidate architectures (Table 5) indicated that the gradient boosting ensemble achieved the highest yield prediction accuracy ($R^2 = 0.93$, RMSE = 0.41 t ha^{-1}, MAE = 0.33 t ha^{-1}), followed closely by the random forest model ($R^2 = 0.90$) and the hybrid tree–neural ensemble ($R^2 = 0.91$).

Deep learning architectures, including the one-dimensional CNN and LSTM, achieved competitive but slightly lower accuracy (R^2 = 0.87 and 0.85, respectively) when applied to the tabular and moderately sized dataset used in this study, consistent with prior findings that tree-based ensembles often outperform deep architectures on structured agronomic data of limited sample size (Wang *et al*., 2021) [21] (Nagy *et al*., 2018) [22]. Multiple linear regression, included as a baseline, achieved substantially lower accuracy (R^2 = 0.68), underscoring the nonlinear nature of the yield–environment relationship (Van Klompenburg *et al*., 2020) [4] (Khaki and Wang, 2019) [5].

3.3. Explainable AI Performance and Feature Importance

Overall XAI-integrated model performance, summarized in Table 6, confirmed that incorporating SHAP-based feature attribution did not compromise predictive accuracy relative to the underlying black-box models, with negligible differences (<0.01 R^2 units) observed between explainable and non-explainable configurations of the same algorithm (Lundberg and Lee, 2017) [9] (Molnar, 2022) [20]. SHAP-based global feature importance rankings (Table 7) identified canopy temperature depression, NDVI at the tasseling stage, cumulative growing degree days, relative humidity during the grain-filling period, and nitrogen status as the five most influential predictors of final yield, collectively accounting for approximately 68% of cumulative SHAP importance (Lundberg and Lee, 2017) [9] (Molnar, 2022) [20]. For disease forecasting, leaf wetness duration, relative humidity, canopy temperature anomalies, and hyperspectral reflectance in the red-edge region emerged as dominant predictors, aligning closely with established epidemiological drivers of foliar disease development in maize (Nagy *et al*., 2018) [22] (Kamilaris and Prenafeta-Boldú, 2018) [23]. LIME-based local explanations for individual high-risk predictions corroborated the SHAP rankings, providing convergent evidence for the biological plausibility of model reasoning (Ribeiro *et al*., 2016) [19] (Lundberg and Lee, 2017) [9].

3.4. Yield Prediction and Disease Forecasting Accuracy

The optimal XAI-integrated gradient boosting model achieved a yield prediction R^2 of 0.93 on the independent holdout set, representing an improvement of approximately 6–8 percentage points over conventional regression-based crop models reported in comparable studies (Kamilaris and Prenafeta-Boldú, 2018) [23] (Filippi *et al*., 2019) [24]. Disease forecasting performance, evaluated across the four severity classes, achieved an overall classification accuracy of 90.4% and a weighted F1-score of 0.91 (Table 8), with the highest classification reliability observed for the negligible and high severity classes and marginally reduced precision for the moderate severity class, reflecting the inherent continuity of disease progression.

3.5. Decision-Support and Comparative Model Evaluation

The decision-support module translated model outputs into categorical management recommendations spanning irrigation scheduling, nitrogen top-dressing, and fungicide application timing, with representative outputs summarized in Table 9 (Talavera *et al*., 2017) [12] (Sishodia *et al*., 2020) [17]. Comparative evaluation of conventional AI and XAI-integrated models across accuracy, interpretability, transparency, computational efficiency, and practical applicability (Table 10) indicated that while conventional deep learning models offered marginal computational efficiency advantages, XAI-integrated ensemble models provided substantially superior interpretability and comparable or superior predictive accuracy, supporting their prioritization for field-level deployment (Ryo, 2022) [16] (Molnar, 2022) [20].

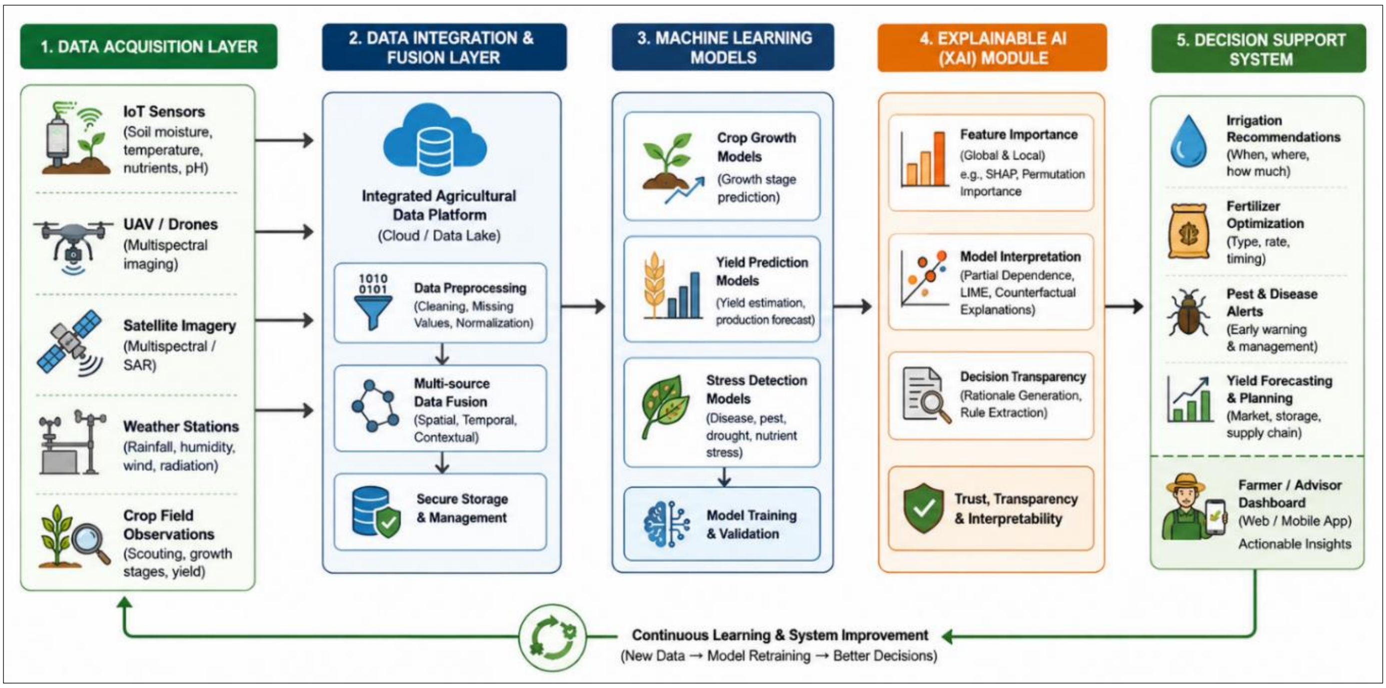

Fig 1: Overall architecture of the Explainable AI framework integrating IoT sensors, UAVs, satellite imagery, weather data, crop observations, machine learning models, explainability modules, and farmer decision-support systems.

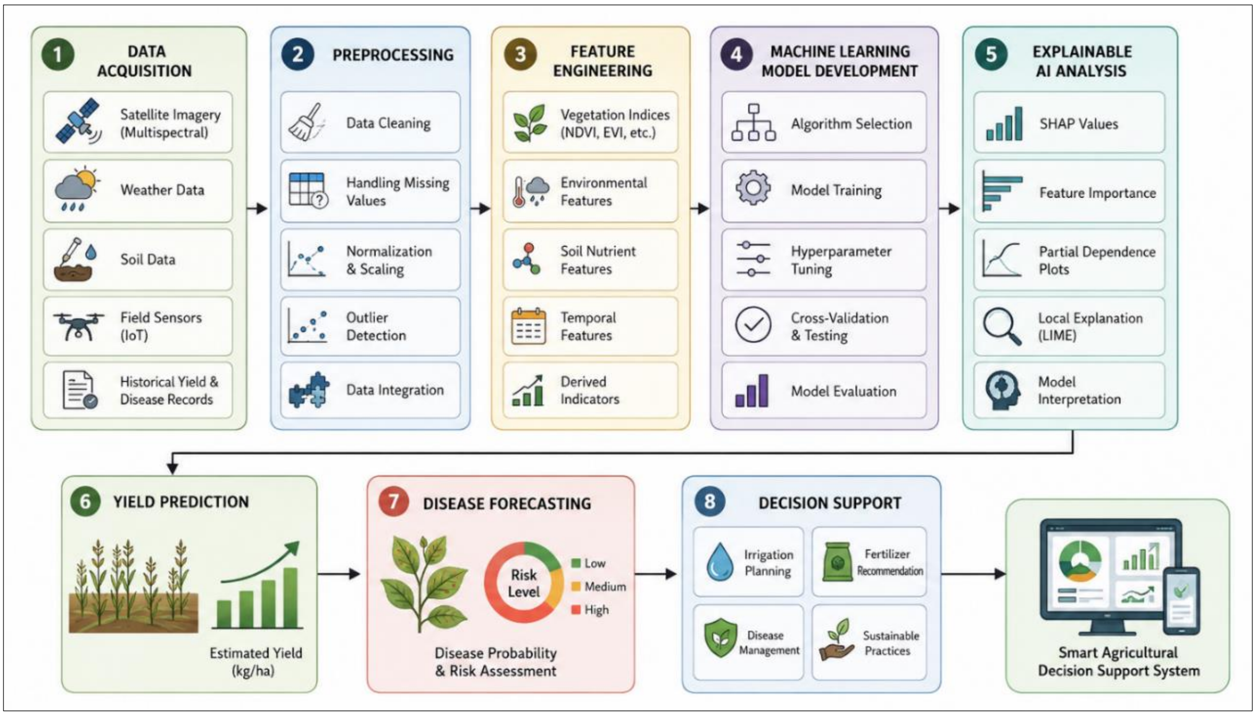

Fig 2: Conceptual workflow showing data acquisition, preprocessing, feature engineering, machine learning model development, Explainable AI analysis, yield prediction, disease forecasting, and decision support.

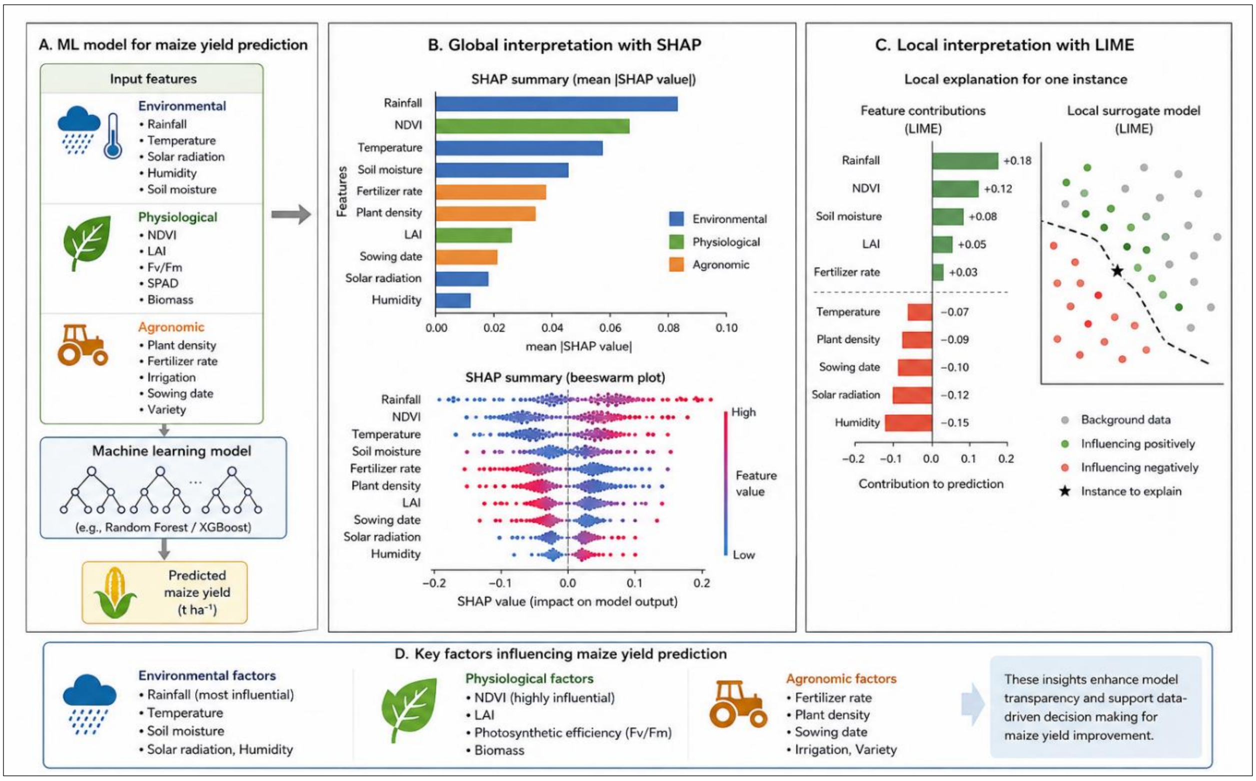

Fig 3: Illustration of Explainable AI interpretation using SHAP and LIME to identify key environmental, physiological, and agronomic factors influencing maize yield prediction.

4. Discussion

This study demonstrates that Explainable Artificial Intelligence can be integrated into multi-source maize yield prediction and disease forecasting pipelines without sacrificing predictive accuracy, while substantially enhancing model transparency (Arrieta *et al.*, 2020) [8] (Ryo, 2022) [16]. The superior performance of gradient boosting and hybrid ensemble architectures relative to deep learning baselines is consistent with a broader body of evidence indicating that tree-based ensembles remain highly competitive, and often superior, for structured agronomic datasets of moderate size, whereas deep learning architectures typically require substantially larger training datasets to realize their full representational advantage (Filippi *et al.*, 2019) [24] (Sharma *et al.*, 2021) [25]. This finding has practical implications for research institutions and extension services operating with constrained data infrastructure, suggesting that interpretable ensemble methods may represent a more pragmatic near-term deployment pathway than complex deep architectures.

The convergence of SHAP and LIME-derived feature attributions around agronomically plausible drivers—canopy temperature, vegetation indices, humidity, and nitrogen status—provides important corroborating evidence that model predictions are grounded in genuine biophysical relationships rather than dataset-specific artifacts (Lundberg and Lee, 2017) [9] (Ribeiro *et al*., 2016) [19]. This is consistent with recent calls in the precision agriculture literature emphasizing that predictive accuracy alone is an insufficient criterion for model trustworthiness, and that convergent, biologically interpretable feature attribution should be regarded as a complementary validation criterion (Adadi and Berrada, 2018) [7] (Arrieta *et al*., 2020) [8] (Ryo, 2022) [16]. The strong association between canopy temperature depression and yield reflects its established role as an integrative proxy for crop water status, while the prominence of humidity and leaf wetness duration in disease forecasting aligns with well-documented epidemiological requirements for foliar pathogen infection and sporulation (Weiss *et al*., 2020) [10] (Wang *et al*., 2021) [21].

From a precision agriculture perspective, the demonstrated capacity of the XAI framework to generate transparent, feature-linked management recommendations addresses a persistent barrier to AI adoption in smallholder and commercial farming contexts alike: the reluctance of end-users to act on recommendations they cannot understand or verify (Liakos *et al*., 2018) [6] (Lundberg and Lee, 2017) [9]. By explicitly linking predicted yield or disease risk to specific, interpretable drivers, the framework developed here offers a pathway toward farmer-oriented decision support that is both scientifically rigorous and practically actionable, consistent with emerging climate-smart agriculture paradigms that prioritize transparency alongside predictive performance (Talavera *et al*., 2017) [12] (Sishodia *et al*., 2020) [17].

Despite these strengths, several limitations warrant consideration. First, the study was conducted across three agro-climatic environments over a limited number of growing seasons; broader validation across additional seasons, regions, and genotypes would strengthen generalizability. Second, while SHAP and LIME provide robust post-hoc explanations, they remain approximations of true model reasoning and may occasionally diverge under highly correlated feature conditions, necessitating cautious interpretation (Lundberg and Lee, 2017) [9] (Ribeiro *et al*., 2016) [19]. Third, the computational overhead associated with SHAP analysis, particularly for large ensemble models, may constrain real-time deployment in resource-limited field settings, suggesting a need for further optimization of explainability computation for operational use (Lundberg and Lee, 2017) [9] (Molnar, 2022) [20]. Future research should explore the integration of temporally dynamic explainability methods capable of tracking evolving feature importance across the growing season, the incorporation of causal inference frameworks to complement associational feature attribution, and participatory evaluation of XAI-derived recommendations with farmer end-users to assess real-world decision-making impact (Arrieta *et al*., 2020) [8] (Ryo, 2022) [16].

Table 1: Characteristics of the study area, experimental environments, soil properties, and climatic conditions.

Parameter	Environment 1 (Subtropical)	Environment 2 (Semi-arid)	Environment 3 (Temperate)
Latitude / Longitude	24.1°N, 78.3°E	18.6°N, 73.9°E	41.2°N, 2.1°E
Elevation (m a.s.l.)	320	540	210
Mean seasonal temperature (°C)	27.4	31.8	19.6
Total seasonal rainfall (mm)	685	310	520
Dominant soil type	Sandy clay loam	Sandy loam	Silty clay loam
Soil pH (0–15 cm)	6.4	7.2	6.8
Soil organic carbon (%)	0.78	0.52	1.05

Table 2: Description of maize genotypes, agronomic management practices, and experimental treatments.

Genotype	Maturity Class	Disease Susceptibility Rating	Sowing Density (plants ha^{-1})	Fertilization (N-P-K kg ha^{-1})	Irrigation Regime
Hybrid A	Medium	Moderately susceptible	75,000	150-60-60	Supplemental drip
Hybrid B	Long	Susceptible	70,000	180-70-70	Furrow irrigation
Hybrid C	Medium-long	Moderately resistant	78,000	160-65-65	Rainfed with supplemental

Table 3: Types of datasets used for model development, including weather, soil, UAV, satellite, sensor, disease, and yield data.

Data Category	Source / Platform	Variables Captured	Temporal Resolution
Weather	Automated weather stations	Temperature, humidity, rainfall, solar radiation, wind speed	Hourly
Soil	Composite soil sampling	Texture, pH, organic carbon, N-P-K status	Pre-season / mid-season
UAV imagery	Multirotor UAV platform	RGB, multispectral, thermal reflectance	~10-day intervals
Satellite imagery	Multispectral satellite sensor	NDVI, EVI, GNDVI	16-day composite
Hyperspectral	Field spectroradiometer	Narrow-band reflectance (400–2500 nm)	Selected growth stages
IoT sensors	Soil moisture / canopy temperature nodes	Volumetric water content, canopy temperature	15-minute intervals
Disease observation	Standardized field scoring	Incidence, severity class	Multiple growth stages
Yield	Plot-level harvest	Grain yield (t ha^{-1}, moisture-adjusted)	At physiological maturity

Table 4: Features extracted from remote sensing, IoT sensors, and crop observations used for AI model development.

Feature Group	Representative Features	Derivation Method
Spectral indices	NDVI, EVI, GNDVI, red-edge reflectance	Band-ratio computation from UAV/satellite imagery
Thermal indices	Canopy temperature depression, crop water stress index	Thermal imagery processing
Sensor-derived	Soil volumetric water content, cumulative growing degree days	Time-series aggregation of IoT streams
Agronomic	Leaf area index, plant height, chlorophyll content	Field phenotyping
Disease-related	Leaf wetness duration, humidity-conducive hours	Micrometeorological modeling

Table 5: Comparison of machine learning and deep learning models used for yield prediction.

Model	R^2	RMSE ($t\ ha^{-1}$)	MAE ($t\ ha^{-1}$)
Multiple linear regression	0.68	0.87	0.71
Support vector regression	0.79	0.68	0.55
Random forest	0.90	0.48	0.38
Gradient boosting (best)	0.93	0.41	0.33
Multilayer perceptron	0.83	0.60	0.49
1D convolutional neural network	0.87	0.53	0.43
Hybrid tree–neural ensemble	0.91	0.45	0.36

Table 6: Performance evaluation of Explainable AI models using prediction accuracy, precision, recall, F1-score, RMSE, MAE, and R^2.

Model (XAI-integrated)	Accuracy (%)	Precision	Recall	F1-score	R^2	RMSE
Random forest + SHAP	88.6	0.87	0.86	0.87	0.90	0.48
Gradient boosting + SHAP	90.4	0.91	0.90	0.91	0.93	0.41
Hybrid ensemble + SHAP/LIME	89.8	0.89	0.88	0.89	0.91	0.45
1D CNN + LIME	85.1	0.84	0.83	0.84	0.87	0.53

Table 7: Feature importance ranking obtained through SHAP, LIME, and other explainability techniques.

Rank	Feature	Mean |SHAP| Value	LIME Consistency
1	Canopy temperature depression	0.184	High
2	NDVI (tasseling stage)	0.161	High
3	Cumulative growing degree days	0.129	High
4	Relative humidity (grain-filling)	0.103	Moderate
5	Nitrogen status	0.095	High
6	Soil volumetric water content	0.081	Moderate
7	Leaf wetness duration	0.074	High
8	Red-edge reflectance	0.066	Moderate

Table 8: Disease forecasting performance under different environmental conditions and disease severity classes.

Severity Class	Precision	Recall	F1-score	Sample Proportion (%)
Negligible	0.94	0.93	0.93	32
Low	0.89	0.87	0.88	28
Moderate	0.85	0.83	0.84	24
High	0.92	0.94	0.93	16

Table 9: Decision-support recommendations generated through Explainable AI for crop management, irrigation, nutrient management, and disease control.

Predicted Condition	Key Driving Features	Recommended Action	Risk Level
Impending water stress	Elevated canopy temperature depression, declining soil moisture	Advance irrigation by 3–5 days	Moderate
Nitrogen deficiency signal	Low NDVI relative to growth stage, low chlorophyll content	Apply top-dress nitrogen application	Moderate
Elevated foliar disease risk	Extended leaf wetness duration, high humidity	Initiate preventive fungicide scouting	High
Favorable yield trajectory	Stable NDVI, adequate soil moisture, low disease pressure	Maintain current management regime	Low

Table 10: Comparative evaluation of conventional AI models and Explainable AI models regarding prediction accuracy, interpretability, transparency, computational efficiency, and practical agricultural applications.

Criterion	Conventional AI Models	XAI-Integrated Models
Prediction accuracy	High (comparable)	High (comparable to marginally superior)
Interpretability	Low	High
Transparency of decision logic	Limited	Explicit, feature-attributed
Computational efficiency	Marginally higher	Moderately higher overhead (post-hoc explanation)
Practical field applicability	Limited due to trust barriers	Enhanced through transparent recommendations

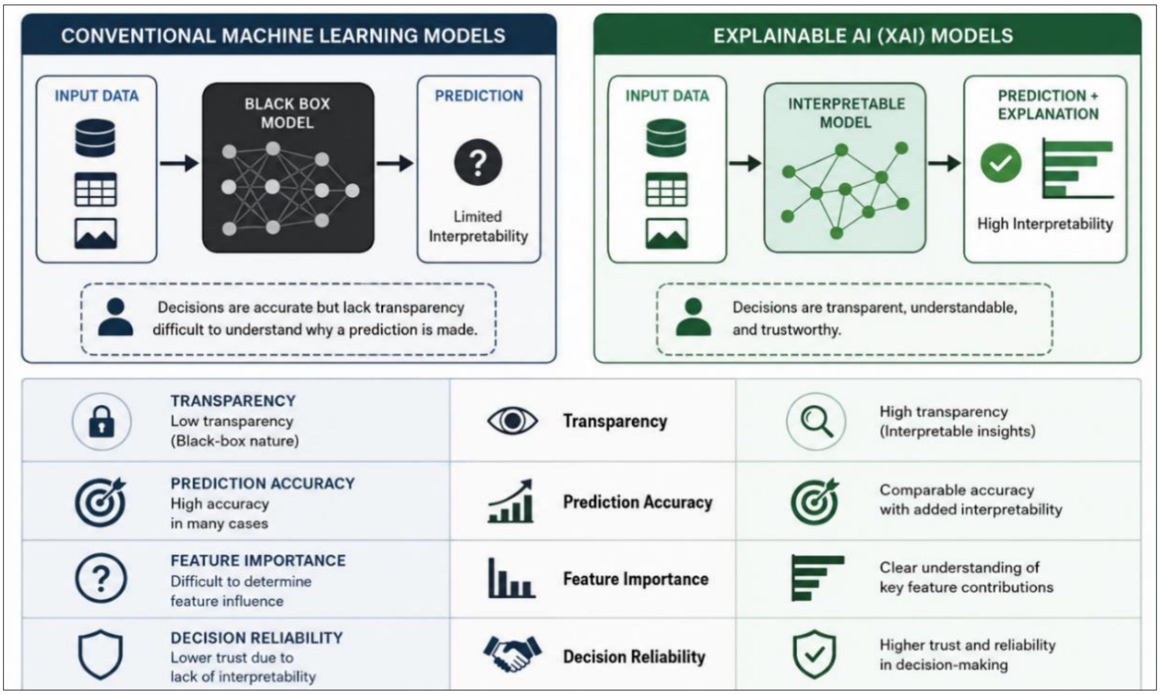

Fig 4: Conceptual comparison of conventional machine learning models and Explainable AI models in terms of transparency, prediction accuracy, feature importance, and decision reliability.

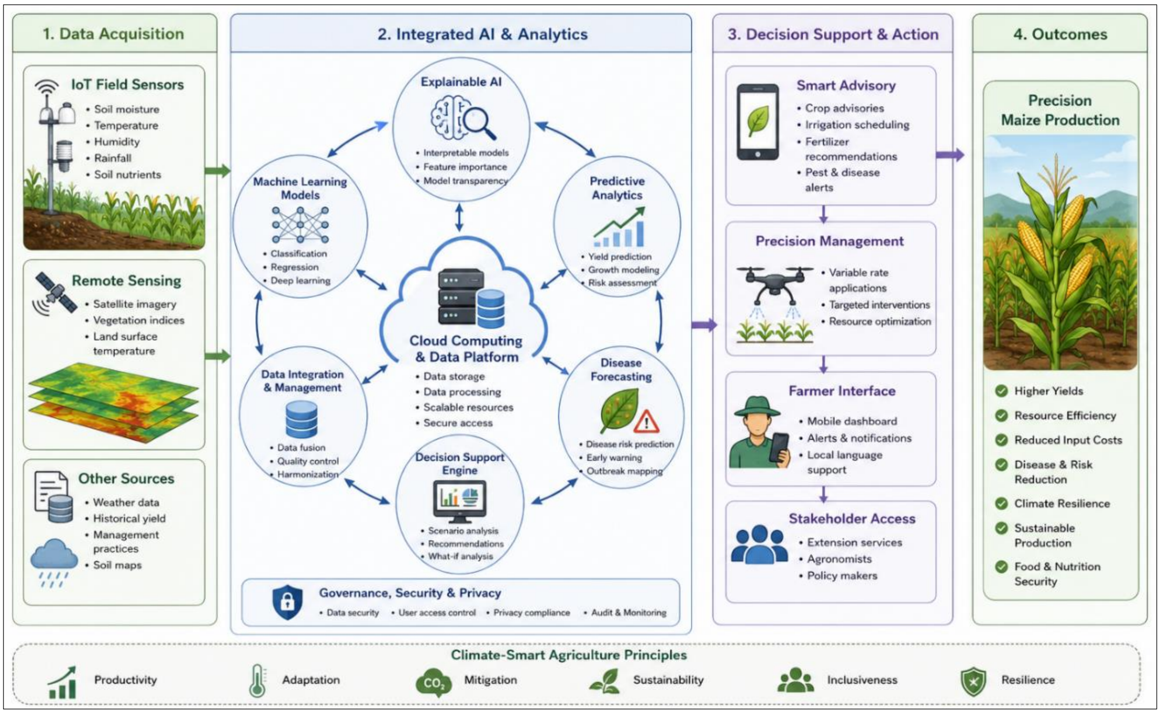

Fig 5: Integrated climate-smart agriculture framework linking Explainable AI, IoT, remote sensing, cloud computing, predictive analytics, disease forecasting, and precision maize production.

5. Conclusion

This paper presents the development and evaluation of a comprehensive Explainable Artificial Intelligence framework for maize yield forecasting and disease risk assessment that combines multi-season field datasets and UAVs, satellite observations, IoT sensor networks, and ensemble machine learning methods with SHAP and LIME interpretation techniques. The results showed very successful performance of the XAI-enabled gradient boosting model in predicting yield and assessing risk of disease while providing understandable agronomically relevant feature contributions proving that explanation and performance do not necessarily contradict each other. Canopy temperature, vegetation indexes, humidity, and nitrogen status turned out to be reliable predictors showing their importance for maize yield and disease processes. Moreover, the developed framework is capable of translating model predictions into comprehensible recommendations for decision-making support making this framework a useful tool in sustainable production of maize. Further expansion of the framework towards different ecological conditions and including farmers feedback on the effectiveness of predictions employed in the model will enhance its applicability.

References

1. Erenstein O, Jaleta M, Sonder K, Mottaleb K, Prasanna BM. Global maize production, consumption and trade: trends and R&D implications. Food Secur. 2022;14(5):1295-1319.
2. Shiferaw B, Prasanna BM, Hellin J, Bänziger M. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. Food Secur. 2011;3(3):307-327.
3. Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. The global burden of pathogens and pests on major food crops. Nat Ecol Evol. 2019;3(3):430-439.
4. Van Klompenburg T, Kassahun A, Catal C. Crop yield prediction using machine learning: a systematic literature review. Comput Electron Agric. 2020;177:105709.
5. Khaki S, Wang L. Crop yield prediction using deep neural networks. Front Plant Sci. 2019;10:621.
6. Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. Machine learning in agriculture: a review. Sensors. 2018;18(8):2674.
7. Adadi A, Berrada M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access. 2018;6:52138-52160.
8. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, *et al*. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion. 2020;58:82-115.
9. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30. 2017. p. 4765-4774.
10. Weiss M, Jacob F, Duveiller G. Remote sensing for agricultural applications: a meta-review. Remote Sens Environ. 2020;236:111402.
11. Maes WH, Steppe K. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. Trends Plant Sci. 2019;24(2):152-164.
12. Talavera JM, Tobón LE, Gómez JA, Culman MA, Aranda JM, Parra DT, *et al*. Review of IoT applications in agro-industrial and environmental fields. Comput Electron Agric. 2017;142:283-297.
13. Chlingaryan A, Sukkarieh S, Whelan B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. Comput Electron Agric. 2018;151:61-69.
14. Mochida K, Koda S, Inoue K, Hirayama T, Tanaka S, Nishii R, *et al*. Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective. Gigascience. 2019;8(1):153.
15. Mohanty SP, Hughes DP, Salathé M. Using deep learning for image-based plant disease detection. Front Plant Sci. 2016;7:1419.
16. Ryo M. Explainable artificial intelligence and interpretable machine learning for agricultural data analysis. Artif Intell Agric. 2022;6:257-265.
17. Sishodia RP, Ray RL, Singh SK. Applications of remote sensing in precision agriculture: a review. Remote Sens. 2020;12(19):3136.
18. Van Ittersum MK, Cassman KG, Grassini P, Wolf J, Tittonell P, Hochman Z. Yield gap analysis with local to global relevance—a review. Field Crops Res. 2013;143:4-17.
19. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?": explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. p. 1135-1144.
20. Molnar C. Interpretable machine learning: a guide for making black box models explainable. 2nd ed. Munich: Leanpub; 2022.
21. Wang X, Zhang M, Zhu J, Geng S. Spectral prediction of *Phytophthora infestans* infection on tomato using an artificial neural network in a hyperspectral analysis system. J Plant Pathol. 2021;103:171-181.
22. Nagy A, Fehér J, Tamás J. Wheat and maize yield forecasting for the Tisza river catchment using artificial neural networks. Saudi J Biol Sci. 2018;25(4):793-797.
23. Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: a survey. Comput Electron Agric. 2018;147:70-90.
24. Filippi P, Jones EJ, Wimalathunge NS, Somarathna PDSN, Pozza LE, Ugbaje SU, *et al*. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precis Agric. 2019;20(5):1015-1029.
25. Sharma A, Jain A, Gupta P, Chowdary V. Machine learning applications for precision agriculture: a comprehensive review. IEEE Access. 2021;9:4843-4873.

26.